

60

*

TGTGGGTTCG AATTTCACCC ACCAACAGTT TTGTCCTACC ATAATTGGGA TAAGGAGTCT

120

*

AATTTCCCTT GTACAATT TTCAATTCTT CCTCCGCCAC ACCATAATA TACTGTACGC

180

*

CACTTCGAAC GCTACAATGT TTGAAAAAG ACGCAGATT TACAAAGACG GAGAAGATAA

TAAGGTTCAA GTACTCCGAT CGTCAGGTGG CCTTTGGAAG CCAACAAACT GGCT ATG

240 *

GCT CTT TCT CTT CAC ACT GTA TTT CTC TGC AAA GAG GAA GCC CTC AAT

Ala Leu Ser Leu His Thr Val Phe Leu Cys Lys Glu Ala Leu Asn

300 *

TTA TAT GCA AGA TCA CCA TGT AAT GAA AGG TTT CAC AGG AGT GGA CAA

Leu Tyr Ala Arg Ser Pro Cys Asn Glu Arg Phe His Asn Ser Gly Gln

360

*

CCT CCT ACC AAC ATA ATC ATG ATG AAA ATT CGA TCC AAC AAT GGA TAT

Pro Pro Thr Asn Ile Ile Met Met Lys Ile Arg Ser Asn Asn Gly Tyr

420

*

TTT AAT TCT TTC CGG TTG TTT ACA TCT TAT AAG ACA AGT TCT TTC TCA
Phe Asn Ser Phe Arg Leu Phe Thr Ser Tyr Lys Thr Ser Ser Phe Ser

FIGURE 1
1/6

GAT	TCT	AGC	CAT	TGC	AAG	GAT	AAA	TCT	CAG	ATA	TGC	AGC	ATC	GAT	ACA
Asp	Ser	Ser	His	Cys	Lys	Asp	Lys	Ser	Gln	Ile	Cys	Ser	Ile	Asp	Thr
480															
*															
AGT	TTT	GAG	GAA	ATA	CAA	AGA	TTT	GAT	CTC	AAA	AGG	GGC	ATG	ACT	TTG
Ser	Phe	Glu	Glu	Ile	Gln	Arg	Phe	Asp	Leu	Lys	Arg	Gly	Met	Thr	Leu
540															
*															
ATT	CTT	GAA	AAG	CAA	TGG	AGA	CAA	TTC	ATA	CAA	TTG	GCT	ATC	GTA	TG
Ile	Leu	Glu	Lys	Gln	Trp	Arg	Gln	Phe	Ile	Gln	Leu	Ala	Ile	Val	Leu
600															
*															
GTT	TGC	ACA	TTT	GTT	ATC	GTT	CCC	AGA	GTT	GAT	GCC	GTT	GAT	GCT	CTT
Val	Cys	Thr	Phe	Val	Ile	Val	Pro	Arg	Val	Asp	Ala	Val	Asp	Ala	Leu
660															
*															
AAA	ACT	TGT	GCT	TGT	TTA	CTC	AAA	GAA	TGC	AGG	ATT	GAG	CTT	GCA	AAA
Lys	Thr	Cys	Ala	Cys	Leu	Leu	Lys	Glu	Cys	Arg	Ile	Glu	Ile	Leu	Ala
720															
*															
TGT	ATA	GCA	AAC	CCA	TCT	TGT	GCG	GCA	AAC	GTT	GCC	TGT	CTA	CAG	ACT
Cys	Ile	Ala	Asn	Pro	Ser	Cys	Ala	Ala	Asn	Val	Ala	Cys	Ile	Cys	Leu
780															
*															
TGC	AAC	AAT	CGT	CCT	GAC	GAG	ACC	GAA	TGT	CAG	ATA	AAA	TGT	GGT	GAC
Cys	Asn	Asn	Arg	Pro	Asp	Glu	Thr	Glu	Cys	Gln	Ile	Lys	Cys	Gly	Asp

FIGURE 1
2/6

780	TTC GAA AAC AGT GTG GAC CAA TTC AAC GAG TGT GCG GTT TCC Leu Phe Glu Asn Ser Val Val Asp Gln Phe Asn Glu Cys Ala Val Ser
840	AAG AAA TGT GTG CCC CGG AAA TCG GAT GTG GGT GAA TTC CCG GTT Arg Lys Cys Val Pro Arg Lys Ser Asp Val Gly Glu Phe Pro Val
900	GAT CGT AAT GCA GTT CAA AAT TTT AAC ATG AAA GAC TTT AGT Pro Asp Arg Asn Ala Val Val Gln Asn Phe Asn Met Lys Asp Phe Ser
960	AAG TGG TAT ATA ACA AGT GGT TTA AAT CCT ACA TTT GAT GCA TTT Gly Lys Trp Tyr Ile Thr Ser Gly Leu Asn Pro Thr Phe Asp Ala Phe
1020	TGT CAA CTT CAT GAG TTT CAT ATG GAA AAT GAT AAA CTT GTT GGG Asp Cys Gln Leu His Glu Phe His Met Glu Asn Asp Lys Leu Val Gly
1080	TTA ACA TGG CGC ATA AAA ACT TTG GAT GGT TTC TTT ACT CGA Asn Leu Thr Trp Arg Ile Lys Thr Leu Asp Gly Gly Phe Phe Thr Arg
	TCT GCT GTG CAA ACA TTT GTT CAA GAT CCA GAT CTT CCT GGA GCA CTT Ser Ala Val Gln Thr Phe Val Gln Asp Pro Asp Leu Pro Gly Ala Leu

FIGURE 1
3/6

TAT	AAT	CAT	GAC	AAT	GAG	TTT	CTT	CAC	CAA	GAT	GAC	TGG	TAC	ATA	
Tyr	Asn	His	Asp	Asn	Glu	Phe	Leu	His	Tyr	Gln	Asp	Asp	Trp	Tyr	Ile
TTA	TCT	TCC	CAA	ATC	GAA	AAA	CCC	GAT	GAT	TAC	ATA	TTC	GTA	TAC	
Leu	Ser	Ser	Gln	Ile	Glu	Asn	Lys	Pro	Asp	Asp	Tyr	Ile	Phe	Val	Tyr
1200	*														
TAC	CGA	GGT	CGA	AAC	GCA	TGG	GAT	GGA	TAC	GGT	GGG	TCC	GTG	ATC	
Tyr	Arg	Gly	Arg	Asn	Asp	Ala	Trp	Asp	Gly	Tyr	Gly	Gly	Ser	Val	Ile
1260	*														
TAC	ACC	CGA	AGC	CCG	ACA	CTC	CCC	GAA	TCG	ATC	ATC	CCA	AAC	CTA	CAA
Tyr	Thr	Arg	Ser	Pro	Thr	Leu	Pro	Glu	Ser	Ile	Ile	Pro	Asn	Leu	Gln
1320	*														
AAA	GCA	GCC	AAA	TCC	GTG	GGT	CGA	GAC	TTT	AAC	AAT	TTC	ATA	ACA	ACC
Lys	Ala	Ala	Lys	Ser	Val	Gly	Arg	Asp	Phe	Asn	Asn	Phe	Ile	Thr	Thr
1380	*														
GAC	AAT	AGT	TGT	GGG	CCT	GAG	CCT	CCA	TTG	GTG	GAA	AGG	CTT	GAG	AAA
Asp	Asn	Ser	Cys	Gly	Pro	Glu	Pro	Pro	Leu	Val	Glu	Arg	Leu	Glu	Lys
ACA	GCG	GAA	GAG	GGC	GAG	AAG	TTG	TTG	ATA	AAA	GAA	GCT	GTA	GAG	ATA
Thr	Ala	Glu	Glu	Gly	Glu	Lys	Leu	Leu	Ile	Lys	Glu	Ala	Val	Glu	Ile

FIGURE 1
4/6

1440	*	GAA GAA GAG GTT GAA AAA GAG GTG GAG AAG GTT AGA GAT ACT GAG ATG
		Glu Glu Glu Val Glu Lys Glu Val Glu Lys Val Arg Asp Thr Glu Met
1500	*	ACT TTG TTT CAG AGG TTG CTT GAA GGG TTT AAG GAG TTG CAA CAA GAT
		Thr Leu Phe Gln Arg Leu Leu Gly Phe Lys Glu Leu Gln Gln Asp
1560	*	GAA GAG AAT TTT GTG AGG GAG TTG AGT AAA GAA GAG AAG GAA ATT CTG
		Glu Glu Asn Phe Val Arg Glu Leu Ser Lys Glu Glu Lys Glu Ile Leu
1620	*	AAT GAA CTT CAA ATG GAA GCG ACT GAA GTT GAA AAG CTT TTT GGG CGC
		Asn Glu Leu Gln Met Glu Ala Thr Glu Val Glu Lys Leu Phe Gly Arg
1680	*	GCG TTA CCG ATT AGG AAA CTT AGA TAAATT CGATGATTGA TTCAGACAAAT
		Ala Leu Pro Ile Arg Lys Leu Arg
1740	*	ATATATAGTC ATATGGATT TGTTAGATACT AGAGAAAAACC CAAAAAAACT TTGTGTATAACG
1800	*	TGATAAACGT GTTTGTGATT TGTTTATTGG CTTAAATAATTG TAGAATAGCT TTTTTAATTG

FIGURE 1
5/6

TTTACAAAAA AATTGATTGT CTATTGGTAG CCAAGAGGTT CACGAAAGA CTGAAAGGGT

1860

*

CTTGGCCGGT TTGCGGGTTA GGCCAAATT TTTGGGGCGG GATCGGTCTT GATCGGGTTT

1920

*

TCTTTAAAAA CATGTATTTT TTATAAATGA TGAGTTATT TCAATTTTTG GCTAAAAAAA

1980

*

1981

*

A

FIGURE 1
6/6

56
TATTTTCATG AGTTTGCACT TGGTGGTAAT ACGGTGAAAG A ATG GCT CTT GCC CCT
His Ser Asn Phe Leu Ala Asn His Glu Thr Ile Lys Tyr Tyr Val Gly
1 5

CAT TCA AAT TTT CTG GCC AAC CAT GAA ACC ATC AAA TAT TAT GTT GGG
His Ser Asn Phe Leu Ala Asn His Glu Thr Ile Lys Tyr Tyr Val Gly
10 15 20

TCA AAG CTG CCC GGT CAT AAA AGG TTT AGC TGG GGT TGG GAA GAT TAC
Ser Lys Leu Pro Gly His Lys Arg Phe Ser Trp Gly Trp Glu Asp Tyr
25 30 35

TTT GGT AGT ATA GTC GTA GCA AAA ATT TGT TCC AGC AGA CGG ATA CCT
Phe Gly Ser Ile Val Ala Lys Ile Cys Ser Ser Arg Arg Ile Pro
40 45 50

AGA TAC TTT CGA AAA TCT CCT AGA ATA TGC TGT GGT TTG GAT TCA AGA
Arg Tyr Phe Arg Lys Ser Pro Arg Ile Cys Cys Gly Leu Asp Ser Arg
55 60 65

GGT CTG CAA CTA TTC TCA CAC GGG AAA CAC AAT CTC TCT CCC GCA CAT
Gly Leu Gln Leu Phe Ser His Gly Lys His Asn Leu Ser Pro Ala His
70 75 80 85

AGC ATT AAC CAG AAT GTA CCT AAG GGA AAT TCA GGA TGC AAA TTT CCA
Ser Ile Asn Gln Asn Val Pro Lys Gly Asn Ser Gly Cys Lys Phe Pro
90 95 100

FIGURE 2
1/5

AAA	GAT	GTA	GCT	TTG	ATG	GTT	TGG	GAG	AAA	TGG	GGC	CAA	TTT	GCC	AAA	392
Lys	Asp	Val	Ala	Ile	Leu	Met	Val	Trp	Glu	Lys	Trp	Gly	Gln	Phe	Ala	Lys
																110
ACA	GCA	ATT	GTA	GCT	ATA	TTC	ATT	TTG	TCA	GTT	GCT	TCA	AAA	GCT	GAT	440
Thr	Ala	Ile	Val	Ala	Ile	Phe	Ile	Leu	Ser	Val	Ala	Ser	Lys	Ala	Asp	120
																125
GCG	GTT	GAT	GCT	CTC	AAG	ACT	TGT	ACT	TGC	TTA	CTG	AAA	GAG	TGC	AGG	488
Ala	Val	Asp	Ala	Leu	Lys	Thr	Cys	Thr	Cys	Leu	Leu	Lys	Glu	Cys	Arg	
																135
																140
																145
TTA	GAG	CTT	GCG	AAG	TGC	ATT	TCG	AAC	CCT	GCA	TGT	GCA	GCT	AAT	GTT	536
Leu	Glu	Leu	Ala	Lys	Cys	Ile	Ser	Asn	Pro	Ala	Cys	Ala	Ala	Asn	Val	
																150
																155
GCC	TGT	CTC	CAG	ACT	TGC	AAC	AAT	AGA	CCT	GAC	GAA	ACG	GAA	TGT	CAG	584
Ala	Cys	Leu	Gln	Thr	Cys	Asn	Asn	Arg	Pro	Asp	Glu	Thr	Glu	Cys	Gln	
																170
																175
ATA	AAA	TGT	GGT	GAT	TTC	TTG	TGG	GAA	AAC	AGT	GTC	GTA	GAG	TTC	AAT	632
Ile	Lys	Cys	Gly	Asp	Leu	Phe	Glu	Asn	Ser	Val	Val	Asp	Glu	Phe	Asn	
																185
																190
																195
GAG	TGT	GCA	GTC	TCC	CGA	AAG	AAA	TGT	GTA	CCT	CGT	AAA	TCT	GAT	GTT	680
Glu	Cys	Ala	Val	Ser	Arg	Lys	Lys	Cys	Val	Pro	Arg	Lys	Ser	Asp	Val	
																200
																205
																210

FIGURE 2
2/5

GGT	GAC	TTT	CCT	GTA	CCT	GAT	CCC	AGT	GTT	CTT	GTC	CAG	AAG	TTC	GAC		728
Gly	Asp	Phe	Pro	Val	Pro	Asp	Pro	Ser	Val	Leu	Val	Gln	Lys	Phe	Asp		
215																225	
ATG	AAA	GAT	TTT	AGC	GGG	AAA	TGG	TTC	ATT	ACT	CGC	GGT	TTG	AAT	CCC		776
Met	Lys	Asp	Phe	Ser	Gly	Lys	Trp	Phe	Ile	Thr	Arg	Gly	Leu	Asn	Pro		
230																240	245
ACT	TTT	GAT	GCT	TTT	GAT	TGC	CAA	TTG	CAT	GAG	TTC	CAT	ACA	GAA	GAA		824
Thr	Phe	Asp	Ala	Phe	Asp	Cys	Gln	Leu	His	Glu	Phe	His	Thr	Glu	Glu		
235																255	260
AAC	AAA	CTT	GTG	GGG	AAT	TTA	TCT	TGG	AGA	ATA	CGT	ACA	CCT	GAT	GGA		872
Asn	Lys	Leu	Val	Gly	Asn	Leu	Ser	Trp	Arg	Ile	Arg	Thr	Pro	Asp	Gly		
240																270	275
GGA	TTT	TTT	ACT	CGA	TCA	GCG	GTG	CAA	AAA	TTC	GTG	CAA	GAT	CCA	AAG		920
Gly	Phe	Phe	Thr	Arg	Ser	Ala	Val	Gln	Lys	Phe	Val	Gln	Asp	Pro	Lys		
245																285	290
TAT	CCG	GGG	ATA	CTC	TAC	AAT	CAT	GAT	AAT	GAG	TAT	CTT	CTC	TAC	CAA		968
Tyr	Pro	Gly	Ile	Leu	Tyr	Asn	His	Asp	Asn	Glu	Tyr	Leu	Tyr	Gln			
250																300	305
GAT	GAC	TGG	TAT	ATT	TTG	TCA	TCC	AAA	GTA	GAA	AAT	AGT	CCA	GAG	GAT		1016
Asp	Asp	Trp	Tyr	Ile	Leu	Ser	Ser	Lys	Val	Glu	Asn	Ser	Pro	Glu	Asp		
255																320	325

FIGURE 2
3/5

TAC	ATA	TTT	GTC	TAC	TAT	AAG	GGC	AGA	AAT	GAT	GCA	TGG	GAT	GGA	TAT	1064	
Tyr	Ile	Phe	Val	Tyr	Tyr	Lys	Gly	Arg	Asn	Asp	Ala	Trp	Asp	Gly	Tyr	340	
																335	
GGT	GGT	TCT	GTA	CTT	TAC	ACA	AGA	AGT	GCA	GTT	TTC	CCT	GAA	AGC	ATT	1112	
Gly	Gly	Ser	Val	Leu	Tyr	Thr	Arg	Ser	Ala	Val	Leu	Pro	Glu	Ser	Ile	350	
																355	
ATA	CCG	GAG	TTG	CAA	ACC	GCA	GCT	CAA	AAA	GTT	GGG	CGT	GAT	TTC	AAC	1160	
Ile	Pro	Glu	Leu	Gln	Thr	Ala	Ala	Gln	Lys	Val	Gly	Arg	Asp	Phe	Asn	360	
																365	
ACA	TTC	ATA	AAA	ACA	GAC	AAT	ACA	TGT	GGC	CCT	GAA	CCT	CCC	CTT	TTT	1208	
Thr	Phe	Ile	Lys	Thr	Asp	Asn	Thr	Cys	Gly	Pro	Glu	Pro	Pro	Leu	Val	375	
																380	
GAG	AGG	TTG	GAG	AAG	AAA	GTG	GAA	GAA	GAA	AGG	ACG	ATC	ATA	AAA	1256	385	
Glu	Arg	Leu	Glu	Lys	Lys	Val	Glu	Glu	Gly	Glu	Arg	Thr	Ile	Ile	Lys	390	
																395	
GAA	GTT	GAG	GAG	ATA	GAA	GAA	GTA	GAG	AAG	GTG	AGA	GAT	AAA	GAA	1304	400	
Glu	Val	Glu	Glu	Ile	Glu	Glu	Glu	Val	Glu	Ilys	Val	Arg	Asp	Lys	Glu	410	
																415	
GTC	ACC	TTA	TTC	AGT	AAA	CTG	TTT	GAA	GGT	TTT	AAA	GAG	CTC	CAA	CGA	1352	420
Val	Thr	Leu	Phe	Ser	Lys	Leu	Phe	Glu	Gly	Phe	Lys	Glu	Ileu	Gln	Arg	425	
																430	
																435	

FIGURE 2
4/5

GAT	GAA	GAG	AAC	TTC	TTA	AGA	GAG	CTG	AGC	AAA	GAA	GAA	ATG	GAT	GTT	1400
Asp	Glu	Glu	Asn	Phe	Leu	Arg	Glu	Leu	Ser	Lys	Glu	Glu	Met	Asp	Val	
																440
																445
																450
																455
																460
																465
																470
																475
																480
																485
																490
																495
																500
																505
																510
																515
																520
																525
																530
																535
																540
																545
																550
																555
																560
																565
																570
																575
																580
																585
																590
																595
																600
																605
																610
																615
																620
																625
																630
																635
																640
																645
																650
																655
																660
																665
																670
																675
																680
																685
																690
																695
																700
																705
																710
																715
																720
																725
																730
																735
																740
																745
																750
																755
																760
																765
																770
																775
																780
																785
																790
																795
																800
																805
																810
																815
																820
																825
																830
																835
																840
																845
																850
																855
																860
																865
																870
																875
																880
																885
																890
																895
																900
																905
																910
																915
																920
																925
																930
																935
																940
																945
																950
																955
																960
																965
																970
																975
																980
																985
																990
																995
																1000

CCACGGGTCC	GGCTTGGTGT	GGGAAAGATT	AGATAGTGTG	AAGA	ATG	GCA	GTA	GCT	56								
ACA	CAT	TGT	TRC	ACT	TCA	CCT	TGT	CAT	GAC	CGT	ATT	CGA	T ^{TTT}	TTC	TCA		1
Thr	His	Cys	Phe	Thr	Ser	Pro	Cys	His	Asp	Arg	Ile	Arg	Phe	Phe	Ser		20
5																	
AGT	GAT	GAT	GGT	ATT	GGG	CTT	GGC	ATT	ACA	AGA	AAG	AGG	ATC	AAT		152	
Ser	Asp	Asp	Gly	Ile	Gly	Arg	Leu	Gly	Ile	Thr	Arg	Lys	Arg	Ile	Asn		35
GGC	ACT	TTC	TRG	CTC	AAG	ATT	TTA	CCT	CCA	ATC	CAA	AGT	GCT	GAT	CTC		200
Gly	Thr	Phe	Leu	Leu	Lys	Ile	Leu	Pro	Pro	Pro	Ile	Gln	Ser	Ala	Asp	Leu	
AGA	ACA	ACT	GGT	GGG	AGA	TCC	TCA	CGT	CCT	TTA	TCT	GCA	TTC	AGG	TCA		248
Arg	Thr	Thr	Gly	Gly	Gly	Arg	Ser	Ser	Arg	Pro	Leu	Ser	Ala	Phe	Arg	Ser	
GGA	TTC	TCT	AAG	GGG	ATA	TTT	GAC	ATT	GTG	CCA	TTA	CCA	TCA	AAG	AAT		296
Gly	Phe	Ser	Lys	Gly	Ile	Phe	Asp	Ile	Val	Pro	Leu	Pro	Leu	Ser	Lys	Asn	
GAG	CTG	AAA	GAG	CTG	ACC	GCT	CCG	CTG	TTG	CTA	AAA	CTC	GTG	GGT	GTT		344
Glu	Leu	Lys	Glu	Leu	Thr	Ala	Pro	Leu	Leu	Leu	Lys	Leu	Leu	Gly	Val		
85																	
90																	
95																	
100																	

FIGURE 3
1/5

TTA	GCT	TGC	GCG	TTC	CTT	ATT	GTT	CCA	TCT	GCA	GAT	GCA	GAT	GCA	392
Leu	Ala	Cys	Ala	Phe	Leu	Ile	Val	Pro	Ser	Ala	Asp	Ala	Val	Asp	Ala
															105
CTT	AAA	ACT	TGT	GCA	TGC	TTA	TTG	AAG	GGA	TGC	AGG	ATA	GAA	CTC	440
Leu	Lys	Thr	Cys	Ala	Cys	Leu	Leu	Lys	Gly	Cys	Arg	Ile	Glu	Leu	Ala
															120
AAG	TGC	ATT	GCC	AAC	CCT	GCC	TGT	GCA	GCC	AAT	GTC	GCG	TGC	CTT	488
Lys	Cys	Ile	Ala	Asn	Pro	Ala	Cys	Ala	Ala	Asn	Val	Ala	Cys	Leu	Gln
															135
ACC	TGC	AAT	AAC	CGT	CCA	GAT	GAA	ACC	GAG	TGC	CAG	ATT	AAA	TGT	536
Thr	Cys	Asn	Asn	Arg	Pro	Asp	Glu	Thr	Glu	Cys	Gln	Ile	Lys	Cys	Gly
															140
Asp	TCG	AGA	AAA	TGT	GTT	CTT	AGT	GTT	GAT	GAG	TTC	AAC	GAG	TGT	584
Leu	Phe	Glu	Asn	Ser	Val	Pro	Arg	Val	Asp	Glu	Phe	Asn	Glu	Cys	Ala
															165
GAT	CTG	TTT	GAG	AAC	AGT	GTT	GAT	GAG	TTC	AAC	GAG	TGT	GCT	CTG	632
Asp	Leu	Phe	Glu	Asn	Ser	Val	Pro	Arg	Lys	Ser	Asp	Leu	Gly	Glu	Pro
															170
TCG	AGA	AAA	AAG	TGT	GTT	CCT	AGA	AAA	TCT	GAT	CTC	GGA	GAA	TTT	680
Ser	Arg	Lys	Lys	Cys	Val	Pro	Arg	Lys	Ser	Asp	Leu	Gly	Glu	Phe	Pro
															185
GCC	CCA	GAC	CCT	TCT	GTT	CTT	GTA	CAG	AAC	TTC	AAC	ATC	TCG	GAC	210
Ala	Pro	Asp	Pro	Ser	Val	Ile	Val	Gln	Asn	Phe	Asn	Ile	Ser	Asp	Phe
															200

FIGURE 3
2/5

bioRxiv preprint doi: <https://doi.org/10.1101/2023.09.11.553212>; this version posted September 11, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a [aCC-BY-ND 4.0 International license](https://creativecommons.org/licenses/by-nd/4.0/).

AAC	GGG	AAG	TGG	TAC	ATT	ACA	AGT	GGC	TTG	AAT	CCA	ACC	TTT	GAT	GCC
Asn	Gly	Lys	Trp	Tyr	Ile	Thr	Ser	Gly	Ileu	Asn	Pro	Thr	Phe	Asp	Ala
215															728
TTC	GAC	TGC	CAG	CTG	CAT	GAG	TTC	CAC	ACA	GAA	GGT	GAC	AAC	AAG	CTT
Phe	Asp	Cys	Gln	Leu	His	Glu	Phe	His	Thr	Glu	Gly	Asp	Asn	Lys	Leu
230															776
GTT	GGA	AAC	ATC	TCT	TGG	AGA	ATA	AAG	ACC	CTA	GAC	AGT	GGA	TTC	TTT
Val	Gly	Asn	Ile	Ser	Trp	Arg	Ile	Lys	Thr	Leu	Asp	Ser	Gly	Phe	Phe
245															824
ACT	AGG	TCA	GCC	GTA	CAA	AAA	TTC	GTG	CAA	GAT	CCT	AAC	CAA	CCT	GGT
Thr	Arg	Ser	Ala	Val	Gln	Lys	Phe	Val	Gln	Asp	Pro	Asn	Gln	Pro	Gly
265															872
GTG	CTC	TAC	AAT	CAT	GAC	AAC	TAC	CAC	TAT	CAA	GAT	GAC	TGG		
Val	Leu	Tyr	Asn	His	Asp	Asn	Glu	Tyr	Leu	His	Tyr	Gln	Asp	Asp	Trp
280															920
TAT	ATC	CTG	TCA	TCA	AAG	ATA	GAG	AAT	AAA	CCT	GAA	GAC	TAT	ATA	TTT
Tyr	Ile	Leu	Ser	Ser	Lys	Ile	Glu	Asn	Lys	Pro	Glu	Asp	Tyr	Ile	Phe
295															968
GTA	TAC	TAC	CGT	GGG	CGA	AAC	GAT	GCT	TGG	GAT	GGA	TAT	GGT	GGT	GCA
Val	Tyr	Tyr	Arg	Gly	Arg	Asn	Asp	Ala	Trp	Asp	Gly	Tyr	Gly	Gly	Ala
310															1016

FIGURE 3
3/5

FIGURE 3
4/5

TTT TTT GAT GAG ATC AAA ATG GAA GCA AGT GAG GTT GAA AAA TTG TTG
Phe Leu Asp Glu Ile Lys Met Glu Ala Ser Glu Val Glu Lys Leu Phe

440 445 450 455 460 465 470 475 480 485 490 495 500
TTT TTT GAT GAG ATC AAA ATG GAA GCA AGT GAG GTT GAA AAA TTG TTG
Phe Leu Asp Glu Ile Lys Met Glu Ala Ser Glu Val Glu Lys Leu Phe

1400 1450 1500 1550 1600 1650 1700 1750 1800 1850 1900 1950 2000
GGG AAA GCT TTG CCA ATC AGG AAG GTC AGG TAGAACAAAG AACCCACATT
Gly Lys Ala Leu Pro Ile Arg Lys Val Arg

440 445 450 455 460 465 470 475 480 485 490 495 500
GTTGTACAAA CTATATTATA CATACTGTGT TCGGTTCATATA TAAAGTAATA TTTTTGTACA
CAGTCATCAT CATTCCATAA CAATTGGATA AAAAAAAA AAAAAA

FIGURE 3
5/5

Tobacco	MALAPHSNFLANHETIKYYVGSKLPGHKRFSWGWDYFGSIVVAKICSSR	50
Arabidopsis	M-V-T.....CFT-PCHDRI--FSS.D-GI-RLGITRK....	33
Lettuce	M--SL-TV--CKE-ALNL-AR-PCNE..--HRS.GQPPTN-IMM--....	43
Tobacco	RIPRYFRKSPRICCGLDLDSRGLQLF.SHGKHNLSPAHSINQNVPKGNSGCK	99
Arabidopsis	--NGT-L..LK-LPPIQ-AD-RTTGRSSRP--AFR-GFSKGIFDIVPLP	81
Lettuce	-SNNGYFN-F-LFTSYKTSSF..SD-SHCKDK-QI.CSIDTSFEEIQRFD	90
Tobacco	FPKDVALMVWEKGQFAKTAIVAIIFILSVASKADA	134
Arabidopsis	SKNELKELTA...PLL-LVG-LACAFLIVPS---	113
Lettuce	LKRGMT-ILEKQ-R--IQL--LVCTFVIVPRV--	125
Tobacco	VDALKTCTCLLKECRIELAKC1SNPACAANVACLQTCCNNRPDETECQIKC	50
Arabidopsis	-----A---G--I-----A-----	50
Lettuce	-----A-----I-----A--S-----	50
Tobacco	GDLFENSVVDEFNECAVSRRKKCVPRKSDVGDFPVDPDSVLVQKFDMKDFS	100
Arabidopsis	-----L-E--A-----N-NIS--N	100
Lettuce	-----Q-----E----RNAV--N-N----	100
Tobacco	GKWFITRGI.NPTFDAFDCLHEFIITE.ENKLVGNLSWRIRTPDGGFFTRS	149
Arabidopsis	---Y--S-----GD-----I-----K-L-S-----	150
Lettuce	---Y--S-----M-ND.-----T---K-L-----	149
Tobacco	AVQKFVQDPKYPGILYNHDNEYLLYQDDWYILSSKVENSPEDYIFVYYKG	199
Arabidopsis	-----NO--V-----H-----I-----K-----R-	200
Lettuce	---T-----DL--A-----F-H-----QI--K-D-----R-	199
Tobacco	RNDAWDGYGGSVLYTRSAVLPESIIPELQTAQKVGRDFNTFIKTDNTCG	249
Arabidopsis	-----A-V---S---N-----EK--KSI---S---R-----	250
Lettuce	-----I-----PT-----N--K--KS-----N--T---S--	249
Tobacco	PEPPIVERLEKKVEEGERTIIKEVEEIEE... <u>EVEK</u> VRDKEVTLFSKLF	295
Arabidopsis	---A---I---T-----I-V----- <u>EVEK</u> -----GRT-M---QR-A	300
Lettuce	-----TA---K.L.--AV----- <u>EVEK</u> -----T-M---QR-L	299
Tobacco	EGFKELQRDEENFLRELSKEEMDVLDGLKMEATEVEKLFGRALPIRKLR	344
Arabidopsis	---N--KQ---V-----EF--EI-----S-----K-----V-	349
Lettuce	-----Q-----V-----KEI-NE-Q-----	348



- A Cysteine-rich domain
- B Lipocalin signature
- C Highly charged domain

FIGURE 4

Percent Identity and Similarity* of Pre-protein VDE

	Lettuce	Tobacco	Arabidopsis
Lettuce		67 (78)	69 (82)
Tobacco	69		68 (81)
Arabidopsis	66	68	

*protein
cDNA*

*similarity values are in parentheses

Percent Identity and Similarity* of Mature VDE

	Lettuce	Tobacco	Arabidopsis
Lettuce		82 (90)	83 (91)
Tobacco	76		83 (92)
Arabidopsis	74	77	

*protein
cDNA*

*similarity values are in parentheses

FIGURE 5

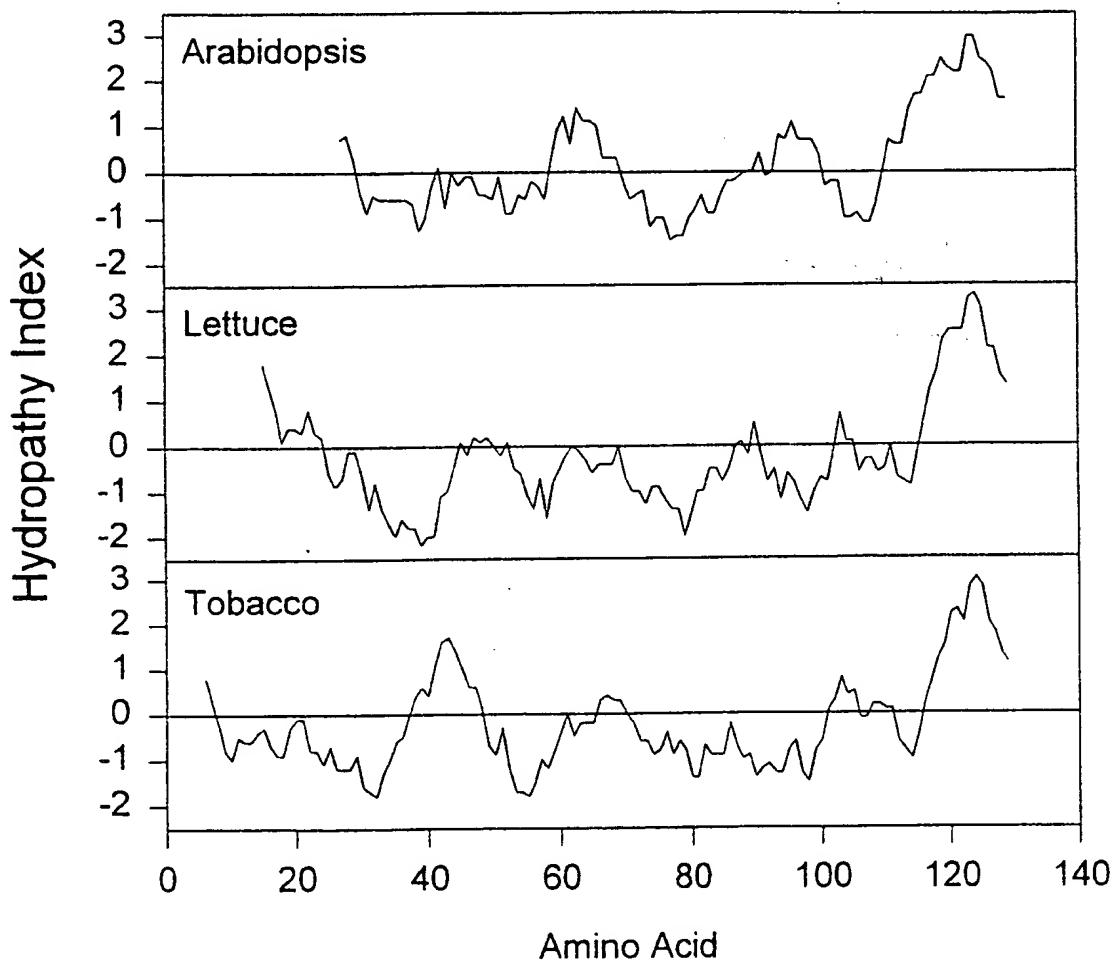


FIGURE 6

100
80
60
40
20
0

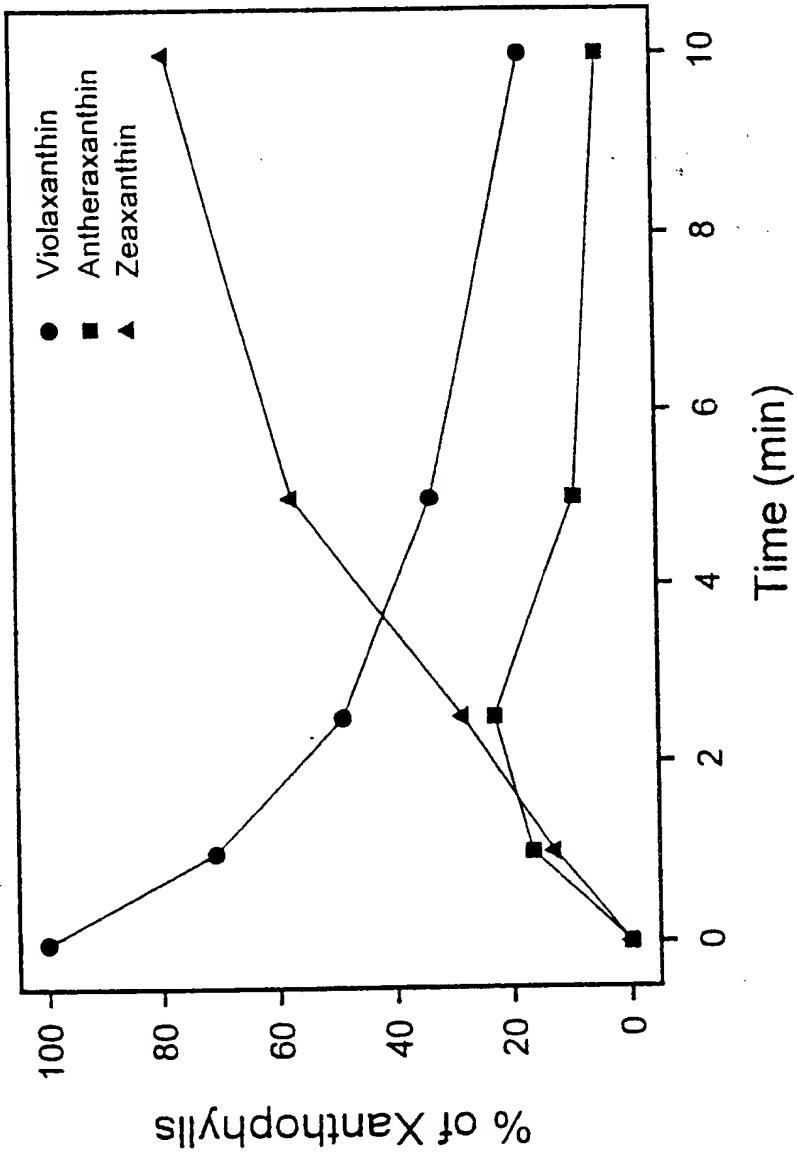


FIGURE 7A

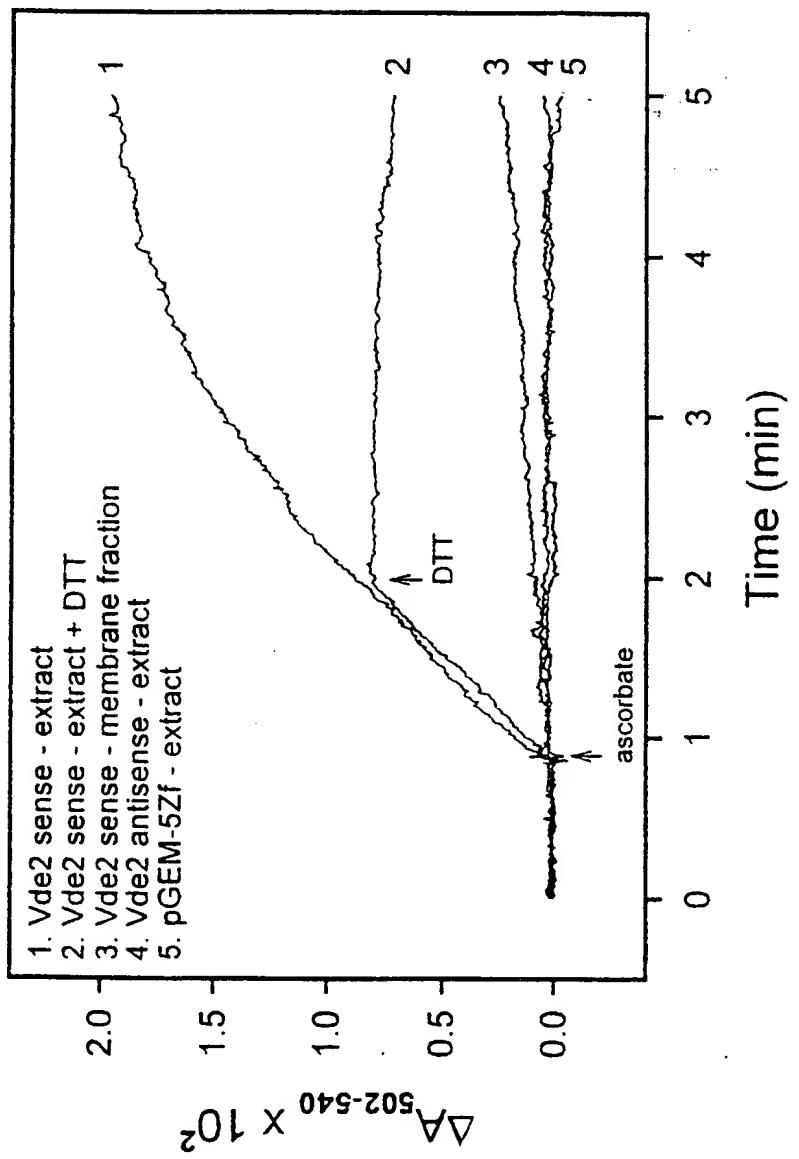


FIGURE 7B

Plant	Treatment	N	V	A	Z	V+A+Z	L	Chl/b/Chla	Bb-Carotene	%V Deepoxidized
Ct-11	Dark	77.13	64.67	1.54	0	66.21	335.12	0.39	136.95	
	Light	77.65	25.56	6.25	30.93	62.74	338.15	0.40	131.76	60.5
Ct-14	Dark	71.60	77.74	1.19	0	78.93	312.05	0.36	150.08	
	Light	72.00	29.07	7.97	43.07	80.11	311.36	0.37	151.50	62.6
Ct-15	Dark	76.68	67.44	0	0	67.44	345.73	0.43	130.05	
	Light	74.45	26.73	7.78	37.44	71.95	337.87	0.42	125.36	60.4
Ct-18	Dark	68.28	82.55	2.33	0	84.88	298.36	0.35	136.67	
	Light	69.65	34.50	13.25	38.44	86.19	311.07	0.36	138.95	58.2
Ct-20	Dark	78.45	70.60	2.85	0	73.45	351.57	0.39	139.58	
	Light	77.38	23.14	5.46	42.66	71.26	343.25	0.39	133.61	67.2
Ct-22	Dark	72.68	104.14	3.40	0	107.54	323.93	0.37	138.29	
	Light	72.13	27.63	6.62	78.66	112.91	315.07	0.40	128.30	73.5
Ct-24	Dark	70.77	76.82	1.55	0	78.37	334.20	0.43	132.95	
	Light	76.52	29.35	7.92	45.24	82.51	339.60	0.44	131.55	61.8
Ct-26	Dark	75.28	63.41	0	0	63.41	346.45	0.44	130.38	
	Light	77.34	26.27	6.16	34.19	66.62	346.91	0.44	128.27	58.6
Ct-30	Dark	78.23	59.66	1.73	0	61.39	357.63	0.45	127.62	
	Light	79.37	26.47	4.93	31.61	63.01	352.39	0.46	124.80	55.6
Ct-31	Dark	71.72	75.91	1.74	0	77.65	315.40	0.37	144.24	
	Light	73.00	31.43	8.74	37.65	77.82	312.80	0.38	145.13	58.6
Ct-39	Dark	75.99	77.93	0	0	77.93	335.79	0.43	127.17	
	Light	74.79	26.28	8.07	41.30	75.65	331.35	0.42	123.11	66.3
Ct-40	Dark	77.56	79.07	2.99	0	82.06	358.33	0.44	126.05	
	Light	77.78	27.44	10.10	47.92	85.46	352.66	0.43	120.89	65.3

Mean = 62.4 ± 5.0

N = 9-cis-neoxanthin V = violaxanthin A = antheraxanthin Z = zeaxanthin Chla = chlorophyll a Chlb = chlorophyll b

All values are relative to chlorophyll a (mmol mol⁻¹ Chla) except Chlb/Chla which is (mol/mol).

FIGURE 8
10
3

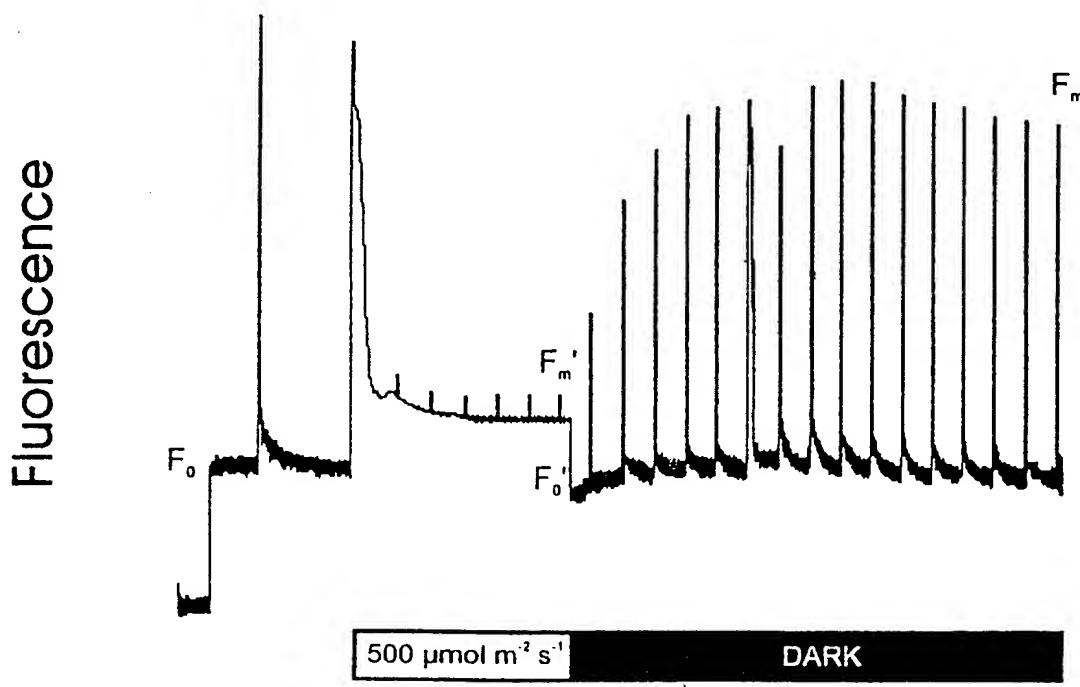
Plant	Treatment	N	V	A	Z	V+A+Z	L	Chlb/Chla	Bb-Carotene	%V De-epoxidized	% Inhibition of De-epoxidation
TAS-32	Dark	74.19	76.98	0	0	76.98	325.75	0.42	136.45	3.7	94.1
	Light	73.78	74.15	2.18	0	76.33	330.95	0.41	131.73		
TAS-39	Dark	77.92	59.19	0	0	59.19	329.29	0.41	141.45	4.7	
	Light	75.06	56.39	2.70	0	59.09	322.29	0.40	141.52		92.5
TAS-21	Dark	75.78	53.19	0	0	53.19	335.21	0.45	132.85		
	Light	77.92	43.90	7.30	9.37	60.57	326.90	0.45	130.33	17.5	72.0
TAS-5	Dark	67.82	79.21	3.43	0	82.64	300.82	0.39	139.00		
	Light	69.72	62.31	14.66	8.27	85.24	300.63	0.40	137.13	21.3	65.9
TAS-17	Dark	74.89	64.54	1.08	0	65.62	317.69	0.41	143.42		
	Light	74.00	49.89	8.49	8.53	66.91	325.32	0.40	139.28	22.7	63.6
TAS-13	Dark	77.92	49.33	1.27	0	50.60	339.63	0.45	135.36		
	Light	78.02	37.82	4.94	7.18	49.94	340.45	0.45	132.78	23.3	62.7
TAS-6	Dark	74.42	55.77	0	0	55.77	340.84	0.44	136.77		
	Light	74.95	40.27	9.69	13.99	63.95	332.00	0.44	135.36	27.8	55.4
TAS-37	Dark	73.05	59.18	1.24	0	60.42	323.30	0.39	135.81		
	Light	71.36	38.97	14.48	9.98	63.43	313.46	0.38	134.62	34.1	45.3
TAS-3	Dark	74.04	60.25	1.76	0	62.01	319.39	0.43	138.89		
	Light	76.98	39.26	7.41	14.33	61.00	322.14	0.44	136.00	34.8	44.2
TAS-36	Dark	69.77	77.86	1.42	0	79.28	295.52	0.36	151.33		
	Light	70.74	48.73	12.76	12.81	74.30	308.06	0.36	151.35	37.4	40.1
TAS-35	Dark	75.59	63.24	1.05	0	64.29	342.09	0.42	130.30		
	Light	75.76	39.48	10.38	17.49	67.35	337.57	0.42	128.88	37.6	39.7
TAS-4	Dark	73.61	68.23	1.31	0	69.54	321.12	0.42	135.43		
	Light	73.23	42.07	8.95	17.84	68.86	320.33	0.42	131.73	38.3	38.6
TAS-9	Dark	72.28	52.57	1.75	0	54.32	324.02	0.42	140.21		
	Light	73.28	31.72	6.19	18.59	56.50	317.11	0.42	136.93	39.7	36.4
TAS-7	Dark	72.55	71.02	1.81	0	72.83	321.37	0.40	133.21		

TAS-38	Light	71.79	39.82	14.04	21.09	74.95	322.04
	Dark	71.66	61.97	1.77	0	63.74	329.67
	Light	73.24	34.45	8.83	19.57	62.85	331.17
TAS-16	Dark	72.15	62.54	2.04	0	64.58	329.72
	Light	74.04	33.28	9.10	23.83	66.21	335.60
TAS-18	Dark	75.09	59.64	1.72	0	61.36	345.04
	Light	75.26	31.68	7.11	23.01	61.80	340.79
TAS-34	Dark	72.35	65.39	1.79	0	67.18	326.06
	Light	72.55	61.22	2.22	0	67.18	326.06

$N = N^{\prime}$ -*cis*-neoxanthin $V =$ violaxanthin $A =$ antheraxanthin $Z =$ zeaxanthin $L =$ lutein Chla = chlorophyll a Chlb = chlorophyll b

All values are relative to chlorophyll a (mol/mol) except Chl b/Chl a which is (mol/mol)

Ct-30

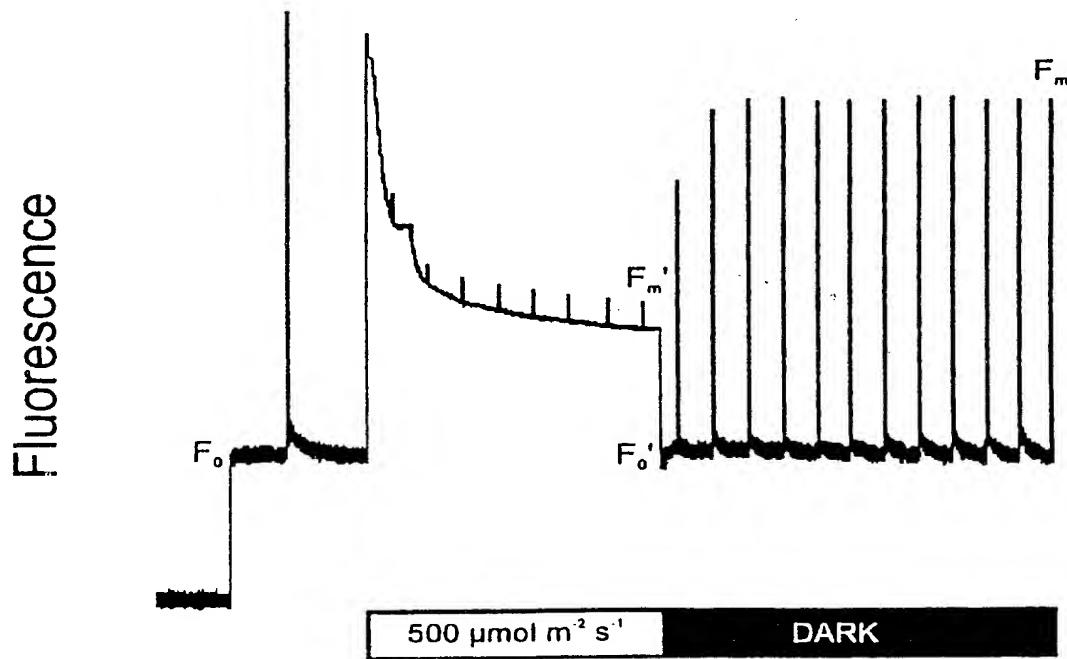


	Dark-adapted	Pre-illuminated	Post-fluorescence Analysis
V	64.28	51.77	44.98
A	1.99	6.16	11.10
Z	0	10.17	13.77
V+A+Z	66.27	68.10	69.85
De-epoxidation (%)		19.5	30.00
$(F_m/F'_m) - 1$			2.20
$(F'_0/F_0) - 1$			0.15

All values are relative to chlorophyll a ($\text{mmol mol}^{-1} \text{ Chla}$).

FIGURE 9

TAS-5



	Dark-adapted	Pre-illuminated	Post-fluorescence Analysis
V	67.51	NA	65.38
A	0	NA	2.14
Z	0	NA	0
V+A+Z	67.51	NA	67.52
De-epoxidation (%)		NA	3.20
$(F_m/F_m') - 1$			1.34
$(F_o/F_o') - 1$			0

All values are relative to chlorophyll a ($\text{mmol mol}^{-1} \text{ Chla}$).

NA - Not assayed

FIGURE 10